

✉ anna.zhukova@pasteur.fr 📍 Institut Pasteur, 25-28 Rue du Dr Roux, 75015 Paris France
🔗 research.pasteur.fr/en/member/anna-zhukova 🔗 linkedin.com/in/annazhukova 🐙 github.com/annazhukova

Research Interests

Phylogenetics: phylodynamics, phylogeography; dynamics of pathogen spreads; mathematical modelling; metabolic modelling; bio-ontologies; algorithms.

Professional Experience

- since 2015 (permanent) *phylogenetics* **Research Engineer, Institut Pasteur, Paris, FR.**
- Modelling drug resistance emergence and transmission in HIV-1
 - Spatio-temporal analysis of virus spread (phylogeography): ZIKV, HIV-1, DENV, Cholera, Sars-CoV-2
 - Epidemiological parameter inference from phylogenies (phylodynamics)
 - Algorithms and tools for large-scale phylogenetics
(Python: *numpy/pandas/scipy/ete3*; JS: *C*; R; *bash*; *Snakemake*; *Docker/Singularity*; *Slurm*).
- 2014–2015 *metabolic modelling* **Postdoc, IBGC/CNRS UMR 5095/U Bordeaux, FR.**
- Constraint-based modelling (FBA, FVA, EFM) of mitochondrial metabolism
(Python: *numpy/pandas/COBRApy*; *SBML*).
- 2011–2014 *metabol. modelling* **Doctoral Researcher, Inria/U Bordeaux/CNRS project-team MAGNOME, Talence, FR.**
- Knowledge-based generalization for metabolic models (Python: *libSBML*; JS: *Leaflet.js*; *SBML*; *ChEBI*; *OBO*).
- 2011 *bio-ontologies* **Intern, EMBL-EBI, Cambridge, UK, Computational Systems Neurobiology group.**
- Development of the Kinetic Simulation Algorithm Ontology (KiSAO) and libKiSAO
(*OWL/OBO*, *OWL API*, *Java*, *HTML*).
- 2010 *misc. bioinfo* **Intern, Inria/U Bordeaux/CNRS joint project-team MAGNOME, Talence, FR.**
- Development of the genome annotation system Magus (*Perl*).
- 2009 *bio-ontologies* **Intern, EMBL-EBI, Cambridge, UK, Functional Genomics group.**
- Experimental Factor Ontology extension of search in the Array Express Archive
(*Java*, *Apache Lucene*, *XML/XSLT*, *OWL API*).
- 2008–2010 *software engineering* **Junior software developer, Swift Teams LLC, St. Petersburg, RU.**
- Amazon Web Services Manager plug-in for IntelliJ IDEA (*Java*, *IntelliJ IDEA*, *AWS API*)
 - JetShot: on-line screen-shot utility for the web-based bug and issue tracker YouTrack (*JetBrains MPS*, *REST*)
 - Actor extension of Java in JetBrains MPS (*JetBrains MPS*, *Java*)

Education

- 2011–2014 **PhD in Computer Science, University of Bordeaux, FR (advisor: DJ Sherman).**
- 2005–2010 **Specialist (BSc+MSc) in Math and Software Engineering, St. Petersburg State University, RU (honours).**
- 2007–2009 **Alumna, Academy of Modern Software Engineering, St. Petersburg, Russia.**

Software and Ontologies [(co-)author]

- PastML** Ancestral scenario reconstruction and visualisation for phylogenies: pastml.pasteur.fr
- PhyloDeep** Deep-learning parameter estimation and model selection from phylogenies: pypi.org/project/phylodeep
- PyBDEI** Birth-Death Exposed-Infectious model parameter estimation from phylogenies: pypi.org/project/pybdei
- Mimoza** Metabolic model visualization and navigation system: mimoza.bordeaux.inria.fr
- MAMO** Mathematical Modelling Ontology: co.mbine.org/standards/mamo
- KiSAO** Kinetic Simulation Algorithm Ontology: co.mbine.org/standards/kisao

Selected Talks [Total: 25 talks, inc. 2 keynotes and 3 invited seminars]

- 20/1/2022 **Fast and Accurate Resolution of the Birth-Death Exposed-Infectious (BDEI) Model, SMPGD**, online.
- 30/6/2020 **A Fast Likelihood Method to Reconstruct and Visualize Ancestral Scenarios, JOBIM**, online.
- 8/9/2018 **Modelling drug resistance spread in HIV, 5th European Student Council Symposium @ ECCB 2018**, Athens Greece, *keynote*.

- 28/11/2016 **A modeling workflow in systems biology: overview**, 1st RSG Luxemb. congress / 2nd Beneluxfr student symposium, Belval Luxembourg, *keynote*.
- 5/8/2015 **Multi-level representation of metabolic networks**, Cheminformatics and Metabolism group seminar, EMBL-EBI, Cambridge UK, *invited seminar*.
- 3/9/2011 **Kinetic Simulation Algorithm Ontology (KiSAO) and libKiSAO**, COMBINE, Heidelberg Germany.

Selected Teaching [Total: 24 courses, 190h]

- 12/2021 **Phylogenetics, Molecular Epidemiology & Phylodynamics**, Montevideo, Uruguay, 5h30.
- 7 – 8/2020 **Molecular evol. & phylogenetics**, Pan African Bioinf. Network for Human Heredity & Health, online, 12h.
- 10/2019 **Math. & Comp. methods for Phylogenetics & Evolutionary Epidemiology**, IPK, Havana Cuba, 7h.
- 10/2018 **Introduction to Molecular Phylogenetics**, HKU, Hong Kong, 7h.
- 3/2017 **Scientific Programming in Python**, IP Paris, France, 5h.
- 3/2015 **Metabolic modelling workshop**, Research sch. Advances in Systems & Syntenic Biol., Strasbourg FR, 2h.
- 2012 – 2013 **Algorithms and programming practical**, University of Bordeaux 1, France, 32h.

Tutoring [Total: 2 PhD students, 5 interns]

- 11/2020 – 11/2021 PhD: **Jakub Voznica** (IP/ENS/CRI, FR), *co-advised the last thesis year with Dr H. Morlon (ENS)*.
 • Deep learning from phylogenies to uncover the transmission dynamics of epidemics
- 11/2020 – 2023(*exp.*) PhD: **Andrew Holtz** (IP/CRI, FR), *co-advising w. Dr H. Bourhy (IP) and Dr G. Baele (KU Leuven)*.
 • Development of ML GLM tool for identification of factors driving viral epidemics
- 6 – 9/2020 intern: **Andrew Holtz** (LSHTM, UK).
 • Development and analysis of novel subsampling protocol for the phylogeography of early SARS-CoV-2 in the US
- 11/19 – 2/2020 interns: **Antoine Verdon** and **Thomas Holer** (CentraleSupélec, FR).
 • Phylogeographic analysis of Cholera spread in ML framework
- 4 – 9/2018 intern: **Luc Blassel** (AgroParisTech, FR), *co-advised with Dr O. Gascuel (IP)*.
 • Machine learning to detect new drug resistance mutations in HIV
- 9 – 12/2017 intern: **Jakub Voznica** (ENS, FR), *co-advised with Dr M. Moslonka and Dr O. Gascuel (IP)*.
 • Impact of partner notification on HIV epidemics control

Grants and scholarships

- 2019 – 2020 **The dynamics of HIV-1 epidemics in Cuba** (Campus France: ~ 10K€).
- 2019 – 2020 **Outcomes of fetuses/infants exposed to Zika virus in utero in Vietnam** (INCEPTION: ~ 10K€).
- 2011 – 2014 **Inria CORDI-S fellowship for doctoral studies** (Inria CORDI-S).

Academic activity

Talk/poster selection for MCEB (2018,2019,2020), AICoB 2020, aSSB 2015
 Invited speaker selection for SARS-CoV-2 session of SMPGD 2022
 Reviewer for Syst. Biol., MBE, Viruses, GBE, Wellcome Open Res., PLOS One, BMC Syst. Biol., etc.
 Reviewer of funding proposals for the University of Oslo (2018), UKRI Medical Research Council (2021).

Languages

Russian (*native*), English (*fluent*), French (*fluent*), Italian (*beginner*)

Other

- 2019, 2021 Among top 3% in **Google Code Jam to I/O for Women** 2019 (30th out of ~ 1000) and 2021 (95th out of ~ 6500).
- 2020-2021 Researcher representative in the Computational Biology Department council of Institut Pasteur.
- 2016,2017,2019 Reviewer of candidates for the Bioinformatics & Biostatistics Hub (Institut Pasteur) IR positions (preselection for the interviews), technical jury member at the interviews in 2017.
- 2013-2014 Inria PhD network organiser; Fr-to-En translator for AFoDIB (Bordeaux comp. sci. PhD student association).

Journal articles

1. Voznica^c, J., Zhukova^c, A., Boskova, V., Saulnier, E., Lemoine, F., Moslonka-Lefebvre, M., & Gascuel^c, O. (2022, July). Deep learning from phylogenies to uncover the epidemiological dynamics of outbreaks. *Nature Communications*, 13(1), 3896. doi: 10.1038/s41467-022-31511-0
2. Shaikh, B., Smith, L. P., Vasilescu, D., Marupilla, G., Wilson, M., Agmon, E., Agnew, H., Andrews, S. S., Anwar, A., Beber, M. E., Bergmann, F. T., Brooks, D., Bruschi, L., Calzone, L., Choi, K., Cooper, J., Detloff, J., Drawert, B., Dumontier, M., Ermentrout, G. B., Faeder, J. R., Freiburger, A. P., Fröhlich, F., Funahashi, A., Garny, A., Gennari, J. H., Gleeson, P., Goelzer, A., Haiman, Z., Hasenauer, J., Hellerstein, J. L., Hermjakob, H., Hoops, S., Ison, J. C., Jahn, D., Jakubowski, H. V., Jordan, R., Kalaš, M., König, M., Liebermeister, W., Sheriff, R. S. M., Mandal, S., McDougal, R., Medley, J. K., Mendes, P., Müller, R., Myers, C. J., Naldi, A., Nguyen, T. V. N., Nickerson, D. P., Olivier, B. G., Patoliya, D., Paulevé, L., Petzold, L. R., Priya, A., Rampadarath, A. K., Rohwer, J. M., Saglam, A. S., Singh, D., Sinha, A., Snoep, J., Sorby, H., Spangler, R., Starruß, J., Thomas, P. J., van Niekerk, D., Weindl, D., Zhang, F., Zhukova, A., Goldberg, A. P., Schaff, J. C., Blinov, M. L., Sauro, H. M., Moraru, I. I., & Karr, J. R. (2022, 05). BioSimulators: a central registry of simulation engines and services for recommending specific tools. *Nucleic Acids Research*. (gkac331) doi: 10.1093/nar/gkac331
3. Seabra^f, S. G., Libin^f, P. J. K., Theys^f, K., Zhukova, A., Potter, B. I., Nebenzahl-Guimaraes, H., Gorbalenya, A. E., Sidorov, I. A., Pimentel, V., Pingarilho, M., de Vasconcelos, A. T. R., Dellicour, S., Khouri, R., Gascuel, O., Vandamme, A.-M., Baele, G., Cuypers, L., & Abecasis, A. B. (2022, 03). Genome-wide diversity of Zika virus: Exploring spatio-temporal dynamics to guide a new nomenclature proposal. *Virus Evolution*, 8(1). (veac029) doi: 10.1093/ve/veac029
4. Fillâtre, P., Dufour, M.-J., Behillil, S., Vatan, R., Reusse, P., Gabelle, A., Velmans, N., Montagne, C., Du Coudret, S. G., Droumaguet, E., Merour, V., Enouf, V., Buzele, R., Valence, M., Guillotel, E., Gagniere, B., Baidaliuk, A., Zhukova, A., Tourdjman, M., Thibault, V., Grohler, C., Pronier, C., Lescure, F.-X., Simon-Loriere, E., Costagliola, D., Van Der Werf, S., Tattevin, P., & Massart, N. (2021, oct). A new SARS-CoV-2 variant poorly detected by RT-PCR on nasopharyngeal samples, with high lethality: an observational study. *Clinical Microbiology and Infection*. doi: 10.1016/j.cmi.2021.09.035
5. Blassel^f, L., Zhukova^f, A., Villabona-Arenas, C. J., Atkins, K. E., Hué, S., & Gascuel, O. (2021). Drug resistance mutations in HIV: new bioinformatics approaches and challenges. *Current Opinion in Virology*, 51, 56–64. doi: 10.1016/j.coviro.2021.09.009
6. Bennedbaek, M., Zhukova, A., Tang, M.-H. E., Bennet, J., Munderi, P., Ruxrungtham, K., Gisslen, M., Worobey, M., Lundgren, J. D., & Marvig, R. L. (2021, jun). Phylogenetic Analysis of HIV-1 Shows Frequent Cross-Country Transmission and Local Population Expansions. *Virus Evolution*. doi: 10.1093/ve/veab055
7. Grant, R., Nguyen, T. T. T., Dao, M. H., Pham, H. T. T., Piorkowski, G., Pham, T. D. T., Cao, T. M., Huynh, L. T. K., Nguyen, Q. H., Vien, L. D. K., Lemoine, F., Zhukova, A., Hoang, D. T. N., Nguyen, H. T., Tran, N. T., Le, L. B., Ngo, M. N. Q., Tran, T. C., Le, N. N. T., Nguyen, M. N., Pham, H. T., Hoang, T. T. D., Dang, T. V., Vu, A. T., Nguyen, Q. N. T., de Lamballerie, X., Pham, Q. D., Luong, Q. C., & Fontanet, A. (2021, jun). Maternal and neonatal outcomes related to Zika virus in pregnant women in Southern Vietnam: An epidemiological and virological prospective analysis. *The Lancet Regional Health - Western Pacific*, 11, 100163. doi: 10.1016/j.lanwpc.2021.100163
8. Zhukova^c, A., Voznica, J., Felipe, M. D., To, T.-H., Pérez, L., Martínez, Y., Pintos, Y., Méndez, M., Gascuel^c, O., & Kouri^c, V. (2021, aug). Cuban history of CRF19 recombinant subtype of HIV-1. *PLOS Pathogens*, 17(8), e1009786. doi: 10.1371/journal.ppat.1009786

9. Zhukova, A., Blassel, L., Lemoine, F., Morel, M., Voznica, J., & Gascuel, O. (2020, nov). Origin, evolution and global spread of SARS-CoV-2. *Comptes Rendus. Biologies*, 0(0), 1–20. doi: 10.5802/crbiol.29
10. Oprea^f, M., Njamkepo^f, E., Cristea^f, D., Zhukova^f, A., Clark, C. G., Kravetz, A. N., Monakhova, E., Ciontea, A. S., Cojocaru, R., Rauzier, J., Damian, M., Gascuel, O., Marie-Laure, Q., & Weill, F.-X. (2020). The seventh pandemic of cholera in Europe revisited by microbial genomics. *Nature Communications*, 11(1), 5347. doi: 10.1038/s41467-020-19185-y
11. Keating, S. M., Waltemath, D., König, M., Zhang, F., Dräger, A., Chaouiya, C., Bergmann, F. T., Finney, A., Gillespie, C. S., Helikar, T., Hoops, S., Malik-Sheriff, R. S., Moodie, S. L., Moraru, I. I., Myers, C. J., Naldi, A., Olivier, B. G., Sahle, S., Schaff, J. C., Smith, L. P., Swat, M. J., Thieffry, D., Watanabe, L., Wilkinson, D. J., Blinov, M. L., Begley, K., Faeder, J. R., Gómez, H. F., Hamm, T. M., Inagaki, Y., Liebermeister, W., Lister, A. L., Lucio, D., Mjolsness, E., Proctor, C. J., Raman, K., Rodriguez, N., Shaffer, C. A., Shapiro, B. E., Stelling, J., Swainston, N., Tanimura, N., Wagner, J., Meier-Schellersheim, M., Sauro, H. M., Palsson, B., Bolouri, H., Kitano, H., Funahashi, A., Hermjakob, H., Doyle, J. C., Hucka, M., & SBML Level 3 Community members. (2020). Sbml level 3: an extensible format for the exchange and reuse of biological models. *Molecular Systems Biology*, 16(8), e9110. doi: 10.15252/msb.20199110
12. Ishikawa^f, S. A., Zhukova^f, A., Iwasaki, W., & Gascuel, O. (2019, may). A fast likelihood method to reconstruct and visualize ancestral scenarios. *Molecular biology and evolution*, 36(9), 2069–2085. doi: 10.1093/molbev/msz131
13. Zhukova, A., Cutino-Moguel, T., Gascuel, O., & Pillay, D. (2017, oct). The Role of Phylogenetics as a Tool to Predict the Spread of Resistance. *The Journal of Infectious Diseases*, 216(suppl_9), S820–S823. doi: 10.1093/infdis/jix411
14. Zhukova, A., Fernandes, L. G., Hugon, P., Pappas, C. J., Sismeiro, O., Coppée, J.-Y., Becavin, C., Malabat, C., Eshghi, A., Zhang, J.-J., Yang, F. X., Picardeau, M., Coppée, J.-Y., Becavin, C., Malabat, C., Eshghi, A., Zhang, J.-J., Yang, F. X., Picardeau, M., & Others. (2017, jan). Genome-wide transcriptional start site mapping and sRNA identification in the pathogen *Leptospira interrogans*. *Frontiers in cellular and infection microbiology*, 7, 10. doi: 10.3389/fcimb.2017.00010
15. Waltemath, D., Karr, J. R., Bergmann, F. T., Chelliah, V., Hucka, M., Krantz, M., Liebermeister, W., Mendes, P., Myers, C. J., Pir, P., Alaybeyoglu, B., Aranganathan, N. K., Baghalian, K., Bittig, A. T., Burke, P. E. P., Cantarelli, M., Chew, Y. H., Costa, R. S., Cursons, J., Czauderna, T., Goldberg, A. P., Gomez, H. F., Hahn, J., Hameri, T., Gardiol, D. F. H., Kazakiewicz, D., Kiselev, I., Knight-Schrijver, V., Knupfer, C., König, M., Lee, D., Lloret-Villas, A., Mandrik, N., Medley, J. K., Moreau, B., Naderi-Meshkin, H., Palaniappan, S. K., Priego-Espinosa, D., Scharm, M., Sharma, M., Smallbone, K., Stanford, N. J., Song, J.-H., Theile, T., Tokic, M., Tomar, N., Toure, V., Uhlenndorf, J., Varusai, T. M., Watanabe, L. H., Wendland, F., Wolfien, M., Yurkovich, J. T., Zhu, Y., Zardilis, A., Zhukova, A., & Schreiber, F. (2016, oct). Toward Community Standards and Software for Whole-Cell Modeling. *IEEE Transactions on Biomedical Engineering*, 63(10), 2007–2014. doi: 10.1109/TBME.2016.2560762
16. Loira, N., Zhukova, A., & Sherman, D. J. (2015, dec). Pantograph: A template-based method for genome-scale metabolic model reconstruction. *Journal of bioinformatics and computational biology*, 13(02), 1550006. doi: 10.1142/S0219720015500067
17. Zhukova^c, A., & Sherman, D. J. (2015, feb). Mimoza: web-based semantic zooming and navigation in metabolic networks. *BMC Systems Biology*, 9(1), 10. doi: 10.1186/s12918-015-0151-5
18. Zhukova^c, A., & Sherman, D. J. (2014a, jul). Knowledge-based generalization of metabolic models. *J Comp Biol*, 21(7), 534–47. doi: 10.1089/cmb.2013.0143
19. Zhukova^c, A., & Sherman, D. J. (2014b, mar). Knowledge-based generalization of metabolic networks: a practical study. *Journal of Bioinformatics and Computational Biology*, 12(2), 1441001. doi: 10.1142/S0219720014410017

20. [Zhukova](#), A., Adams, R., Laibe, C., & Le Novère, N. (2012). LibKiSAO: a Java Library for Querying KiSAO. *BMC Research Notes*, 5(1), 520+. doi: 10.1186/1756-0500-5-520
21. Courtot^f, M., Juty^f, N., Knüpfer^f, C., Waltemath^f, D., [Zhukova](#)^f, A., Dräger, A., Dumontier, M., Finney, A., Golebiewski, M., Hastings, J., Hoops, S., Keating, S., Kell, D. B., Kerrien, S., Lawson, J., Lister, A., Lu, J., Machne, R., Mendes, P., Pocock, M., Rodriguez, N., Villeger, A., Wilkinson, D. J., Wimalaratne, S., Laibe, C., Hucka, M., & Le Novère, N. (2011, jan). Controlled vocabularies and semantics in systems biology. *Molecular systems biology*, 7(1), 543. doi: 10.1038/msb.2011.77
22. Malone, J., Holloway, E., Adamusiak, T., Kapushesky, M., Zheng, J., Kolesnikov, N., [Zhukova](#), A., Brazma, A., & Parkinson, H. (2010). Modeling sample variables with an Experimental Factor Ontology. *Bioinformatics*, 26(8), 1112–1118. doi: 10.1093/bioinformatics/btq099

Conference proceeding articles

23. [Zhukova](#), A., & Mazat, J.-P. (2015). The mitochondrion metabolic model iAS253 revisited. *advances in Systems and Synthetic Biology*, 179–188.
24. [Zhukova](#), A., & Mazin, M. (2010). Actor extension of Java language in MPS Environment. *Sci Tech J Inf Technol Mech Opt*, 10(2), 72–77.

Book chapters

25. [Zhukova](#), A., Gascuel, O., Duchene, S., Ayres, D., Lemey, P., & Baele, G. (2020). Efficiently Analysing Large Viral Data Sets in Computational Phylogenomics. In C. Scornavacca, F. Delsuc, & N. Galtier (Eds.), *Phylogenetics in the genomic era* (pp. 5.3:1–5.3:43). No commercial publisher — Authors open access book.

PhD Thesis

26. [Zhukova](#), A. (2014). *Knowledge-based scaling for biological models* (Doctoral dissertation, Université Bordeaux-I). Retrieved from <http://www.theses.fr/2014BORD0427>

Preprints

27. [Zhukova](#)^c, A., Hecht, F., Maday, Y., & Gascuel^c, O. (2022). Fast and accurate maximum-likelihood estimation of Birth-Death Exposed-Infectious epidemiological model from phylogenetic trees. *medRxiv*, 2022.08.02.22278328. doi: 10.1101/2022.08.02.22278328